

MICROPROCESSOR TECHNOLOGY

Assis. Prof. Hossam El-Din Moustafa

Lecture 3

Ch.1 The Evolution of The Microprocessor

Chapter Objectives

- Introduce the microprocessor evolution from transistors to integrated circuits.
- Describe Moore's law and some of its implications.
- List possibilities for future of Moore's law.

Microprocessor Evolution

- Supercomputers are designed to perform calculations using hundreds or thousands of microprocessors
- Personal computers that have a single central processor use other processors to control the display, network communication, disk drives, and other functions.
- Cars, stereos, cell phones, microwaves, and washing machines all contain microprocessors
- The possible uses of the processor are limited only by the imagination of the programmer.

The Keys to Microprocessor Success

1. Flexibility
2. Steady improvement of performance

People will sometimes say they need to buy a new computer because their old one has become *too slow*. This is of course only a matter of perception. Their computer has the exact same speed as the day they bought it.

What has changed to make it appear slower is the software.

The Transistor

- Despite all the different functions a microprocessor performs, at the end it is only a collection of transistors and wires.
- The job of microprocessor design is ultimately deciding how to connect transistors to be able to quickly execute the commands that run programs.
- The story of the first microprocessor is therefore also the story of the invention of the transistor and the integrated circuit.

The Transistor

- Basic concepts (valence band, Conduction band)
- P-N junction review (1940)
- Forward bias vs. reverse bias
- Bardeen and Brattain → the transistor (1947)
- Shockley → Junction transistor (1948)
- Sparks and Teal → BJT 1950)
- The first commercially available transistors were all made of **germanium**.

The Transistor

- They tended to be very sensitive to temperature and at temperatures above 75°C they didn't work at all.
- The problem was that the band gap of germanium was too small.
- **Silicon's** band gap is almost twice as large as germanium's, so it is far less sensitive to temperature and has a much higher maximum operating temperature.
- In 1954, TI produced the first silicon junction transistor.

The Integrated Circuit

- **Kilby** had constructed the first Integrated Circuit (IC) (1958).
- **Noyce** imagined simultaneously making not only transistors but the wires connecting them as well.

IC (*Cont.*) Making Connections

- Silicon naturally forms silicon dioxide (the main ingredient in glass) when exposed to air and heat, and silicon dioxide is an excellent insulator.
- By growing a layer of silicon dioxide on top of a silicon chip, the components could be isolated from wires deposited on top. Acid could be used to cut holes in the insulator where connections needed to be made.

IC (*Cont.*) Field Effect Transistor

- Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) (1960)
- Where bipolar junction transistors require a constant current into the base to remain switched on, MOSFETs require only a voltage to be held at the gate.
- This allows MOSFETs to use less power than equivalent BJT circuits.
- Almost all transistors made today are MOSFETs connected together in integrated circuits.

IC (*Cont.*) Field Effect Transistor

- The key to making integrated circuits cost effective enough for the general market place was incorporating more transistors into each chip.
- The computers of the 1960s stored their data and instructions in “**core**” memory. These memories were constructed of grids of wires with metal donuts threaded onto each intersection point.
- By applying current to one vertical and one horizontal wire a specific “core” could be magnetized in one direction or the other to store a single bit of information.
- Core memory was reliable but difficult to assemble and operated slowly compared to the transistors performing computations. A memory made out of transistors was possible but would require thousands of transistors to provide enough storage to be useful.

Intel® Integrated electronics

- *Intel (1968) by Noyce and Moore*
 1. First Intel IC, 256 bit memory chip (1101) in (1969)
 2. 1024 bit memory chip (1103) in (1970)
 3. Chips for calculator (1971)
 - 1) A chip controlling input and output functions
 - 2) A memory chip to hold data
 - 3) A chip to hold instructions
 - 4) A central processing unit that would eventually become the world's first microprocessor (4004)
(The 4004 contained 2300 transistors (24mm²) and ran at a clock speed of 740 kHz, executing on average about 60,000 instructions per second).
- The 4004 used transistors with a feature size of 10 microns

Moore's Law

- **Process scaling:** shrinking the physical size of the transistors and the wires interconnecting them, allowing more devices to be placed on each chip, which allows more complex functions to be implemented.
- shrinking transistor dimensions were allowing the number of transistors on a die to double roughly every 18 months (**Moore, 1975**)
- For microprocessors, the trend has been closer to a doubling every 2 years.

Moore's Law

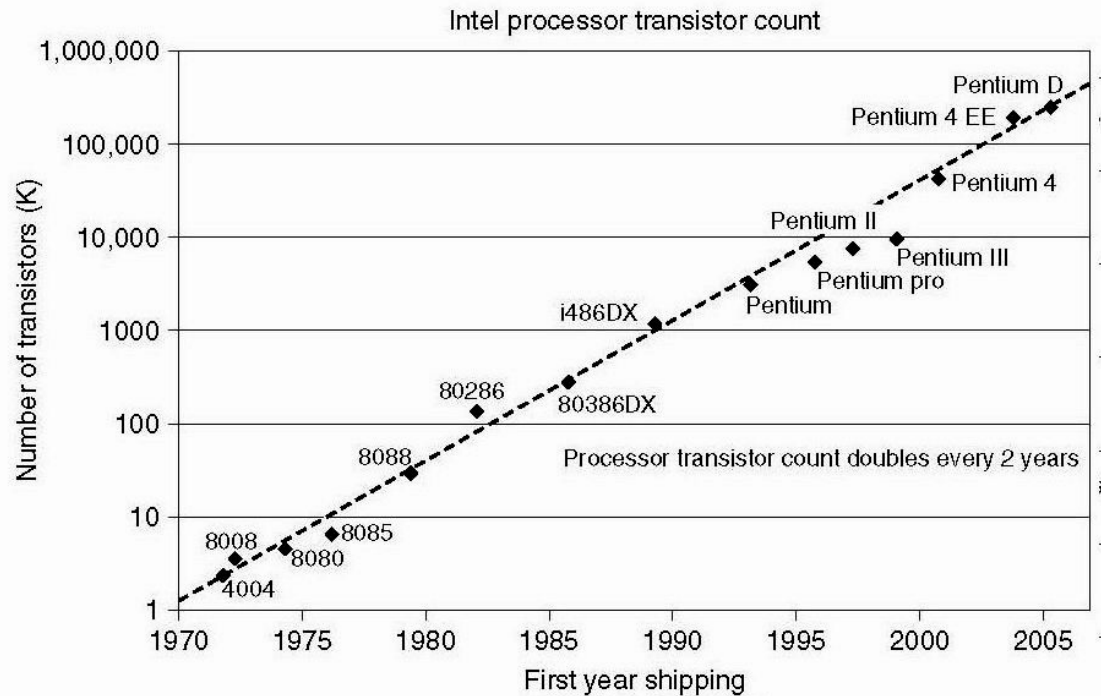


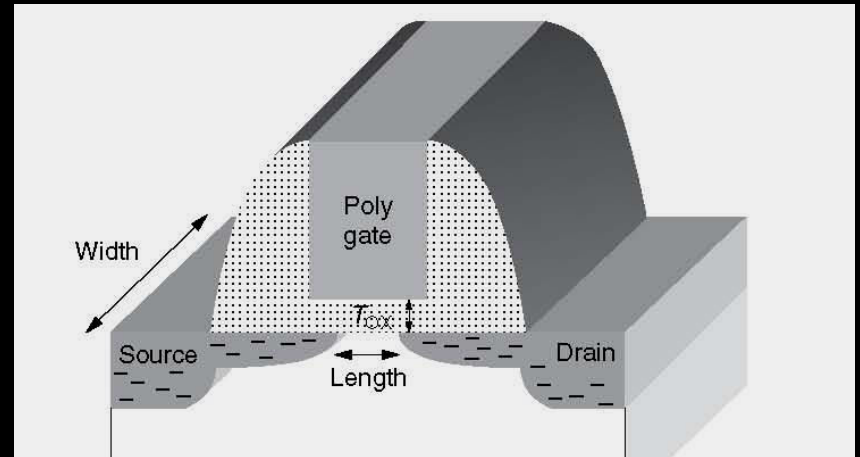
Figure 1-7 Moore's law. (Moore, "No Exponential is Forever," 10.)

Transistor Scaling

- The reason smaller transistors switch faster is that although they draw less current, they also have less capacitance
- The capacitance of the gate increases linearly with the width (W) and length (L) of the gate and decreases linearly with the thickness of the gate oxide.

Transistor Scaling

- Reducing the gate oxide thickness increases current, since pushing the gate physically closer to the silicon channel allows its electric field to better penetrate the semiconductor and draw more charges into the channel.
- Decreasing the device length is the most effective, and this is what the semiconductor industry has focused on the most



Interconnection Scaling

- To connect millions of transistors, modern microprocessors may use seven or more separate layers of wires.
- They add capacitive load to the transistor outputs, and their resistance means that voltages take time to travel their length.
- The capacitance of a wire is the sum of its capacitance to wires on either side and to wires above and below.
- In 2000, some semiconductor manufacturers switched from using aluminum wires, to copper wires (*Why*)?? Why copper was not used?? *See p. 14*

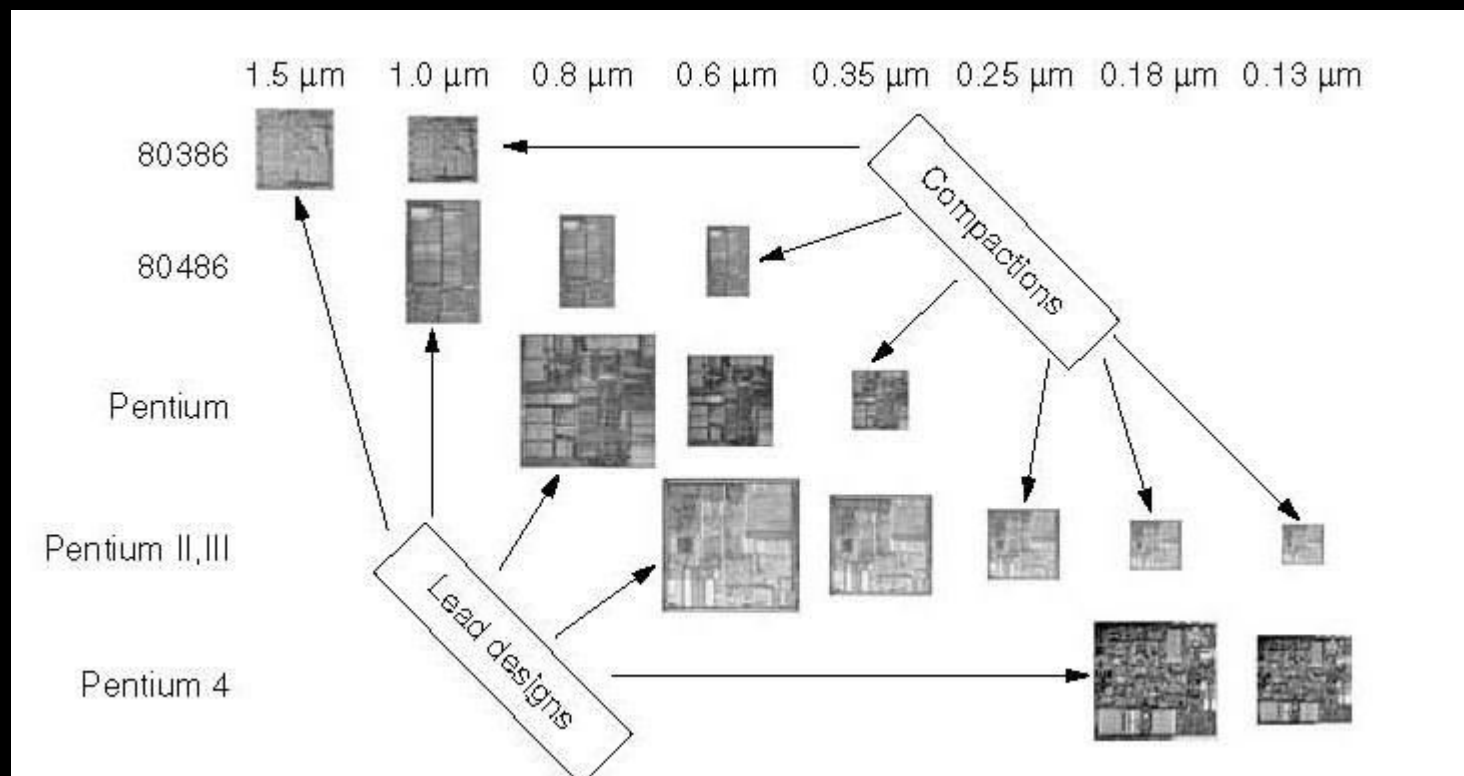
Interconnection Scaling

- Wire capacitances have been reduced through the use of low- K dielectrics.
- Not only the dimensions of the wires determine wire capacitance but also by the permittivity or K value of the insulator surrounding the wires.
- The best capacitance would be achieved if there were simply air or vacuum between wires giving a K equal to 1, but of course this would provide no physical support.
- Silicon dioxide is traditionally used, but this has a K value of 4. New materials are being tried to reduce K to 3 or even 2, but these materials tend to be very soft and porous.

Microprocessor Scaling

- Microprocessor designs can be broken down into two basic categories: *lead designs* and *compactions*.
- *Lead designs* are fundamentally new designs. They typically add new features that require more transistors and therefore a larger die size.
- *Compactions* change completed designs to make them work on new fabrication processes. This allows for higher frequency, lower power, and smaller dies.

Microprocessor Scaling



Microprocessor Scaling

- Semiconductor industry has begun a new generation of fabrication process every 2 to 3 years. Each generation reduces horizontal dimensions about 30 percent compared to the previous generation.
- As an example, the 180-nm generation Intel Pentium® 4 began at a maximum frequency of 1.5 GHz and scaled to 2.0 GHz. The 130-nm Pentium 4 started at 2.0 GHz and scaled to 3.4 GHz. The 90-nm Pentium 4 started at 3.2 GHz.
- Each new technology generation is planned to start when the previous generation can no longer be easily improved.

The Future of Moore's Law

- No exponential trend can continue forever, and this simple fact has led to predictions of *the end* of Moore's law for decades.
- If it is to remain true in the future, it will be because the industry finds it profitable to continue to solve “insurmountable” problems and force Moore's law to come true.
- There have already been a number of new fabrication technologies proposed or put into use that will help continue Moore's law till 2015.

The Future of Moore's Law

- Multiple threshold voltages
- Silicon on insulator (SOI)
- Strained silicon
- High K-gate dielectrics
- Improved interconnections
- Double-triple gate

Multiple Threshold Voltages

- By applying different amounts of dopant to the channels of different transistors, devices with different threshold voltages are made on the same die.
- When speed is required, low V_T devices, which are fast but high power, are used.
- In circuits that do not limit the frequency of the processor, slower, more power efficient, high V_T devices are used to reduce overall leakage power.
- It is used in the Intel 90-nm generation.

Silicon on Insulator (SOI)

- SOI transistors, build MOSFETs out of a thin layer of silicon sitting on top of an insulator.
- This layer of insulation reduces the capacitance of the source and drain regions, improving speed and reducing power.
- SOI is already in use in the Advanced Micro Devices (AMD®) 90-nm generation.

Strained Silicon

- The ability of charge carriers to move through silicon is improved by placing the crystal lattice under strain
- Depositing *silicon nitride* on top of the source and drain regions tends to compress these areas. This pulls the atoms in the channel farther apart and improves *electron mobility*.
- Depositing *germanium atoms*, which are larger than silicon atoms, into the source and drain tends to expand these areas. This pushes the atoms in the channel closer together and improves *hole mobility*.
- It is used in the Intel 90-nm generation.

High K-Gate Dielectric

- Gate oxide layers thinner than 1 nm are only a few molecules thick and would have very large gate leakage currents.
- Replacing the silicon dioxide, which is currently used in gate oxides, with a higher permittivity material strengthens the electric field reaching the channel.
- This allows for thicker gate oxides to provide the same control of the channel at dramatically lower gate leakage currents.

Improved Interconnections

- Improvements in interconnect capacitance are possible through further reductions in the permittivity of inter-level dielectrics
- However, improvements in resistance are probably not possible.
- Improving the scaling of interconnects is currently the greatest challenge to the continuation of Moore's law.

Double-Triple Gate

- Wrap the gate wire around two or three sides of a raised strip of silicon.
- In a triple gate device the channel is like a tunnel with the gate forming both sides and the roof.
- This allows strong electric fields from the gate to penetrate the silicon and increases on current while reducing leakage currents.
- These ideas allow at least an educated guess as to what the scaling of devices may look like over the next 3 years

Microprocessor Fabrication Future

Table (1.1) Microprocessor Fabrication Projection Till 2015

New generation every 2–3 years

30% reduction in gate length

30% increase in gate capacitance through high- K materials

15% reduction in voltage

30% reduction in interconnect horizontal and vertical dimensions for lower metal layers

15% reduction in interconnect horizontal and vertical dimensions for upper metal layers

Add 1 metal layer every generation

Thank You

With all best wishes !!